SUPPLEMENTARY METHODS

Strains and growth conditions

To quantify the fitness impact of evolved mutations, we used a broad-host-range vector (pCM433), for marker-free and scar free allelic exchange (Marx 2008) to move mutations into the ancestral genomic background of *M. extorquens* strains. We first cloned the region containing the evolved mutation into pCM433 using appropriate enzymes (Table S3, set "Cloning/sequencing (*fae* SNPs)"; Table S4); transformed *E. coli* K-12 MG1655 with the plasmid; confirmed the presence of the insert in *E. coli* clones by colony PCR (Table S3, set "sequencing"); and set up tri-parental matings with a recipient strain of *M. extorquens* and a helper *E. coli* strain (bearing the pRK2073 plasmid to facilitate the transfer of pCM433 from the donor cell into the recipient cell). Similarly, to quantify the combined fitness impact of mutations in the rest of the genome (evolved genomic background; set "bkg"), we moved the ancestral *fae* alleles into clones from the respective evolved lines, while replacing the evolved *fae* allele. We set up tri-parental matings between the evolved clone of *M. extorquens* (set "evo"), an *E. coli* K-12 MG1655 strain carrying the appropriate fae allele on the pCM433 plasmid (generated in Agashe *et al*, 2013; set "Cloning") and the pRK2073 helper strain. We confirmed final strains by PCR (Table S3, set "sequencing") and then tested their fitness, *fae* mRNA levels and protein levels.

To generate alleles carrying an N terminal insert sequence enriched in frequently used codons (FLAG tag with additional residues for cloning: ATGGACTACAAGGATGACGATGACAAGGCTGCAGCGCAT), we digested the *de novo*-synthesized *fae* alleles from an M13 plasmid vector (described in (Agashe et al. 2013)) with *Pst*I and *Mlu*I. We cloned the insert into pCM433 containing the *fae* upstream and downstream regions as described previously in (Agashe et al. 2013). We then moved alleles carrying an N terminal insert sequence from pCM433 into the cumate-inducible expression vector, pHC115 (12) (Table S3, set "cloning (N-ter tag)"). We induced *fae* expression from these plasmids using 1 μM cumate (Sigma) in Hypho + 15 mM Methylamine. We determined growth rates for these alleles as described later.

To test the impact of predicted beneficial and detrimental point mutations in ancestral WT, AF, VA and AC alleles, we created mutant alleles using site-directed mutagenesis primers and cloned these into the expression vector pLC291 (Chubiz et al. 2013) using the appropriate enzymes (Table S3, set "cloning (sd)"). To test whether beneficial mutations were advantageous only in the context of the specific *fae* allele in which they arose, we used site-directed mutagenesis primers and cloned them as above (Table S3, set "cloning (cs)"). These constructs were then transformed into *E. coli* K-12 MG1655, and Kanamycin-resistant positive clones (Table S4, set "sd", "cs") were confirmed by colony PCR (Table S3, set "sequencing"). We transformed the *M. extorquens* AM1 *fae* knockout (Δ*fae*; TableS2, set "anc-orig") as described above. Successful transformants were selected in Hypho-succinate agar plates with Kanamycin

and Rifamycin. We induced *fae* expression from plasmids using 100 ng/ml anhydrotetracycline (Sigma) in Hypho + 15 mM Methylamine. We determined growth rates for all ancestral, evolved and mutant alleles as described later.

Experimental evolution

Small population size: Each replicate population of a mutant ancestor was initiated using a single colony picked from a Hypho-succinate agar plate (Table S2, set "anc") and allowed to grow in Hypho-succinate broth for 48 hours before initiating experimental evolution. We inoculated six replicate populations of each mutant in a checkerboard pattern in 48-well microplates (Costar, Corning) maintained at 30 °C in a shaking incubator at 650 rpm. Every 48 hours, we transferred 10 µL growing culture to 630 µL fresh Hypho media (1:64 dilution) with 15 mM methylamine as the sole carbon source. This protocol allows for six generations per transfer in the WT. We frequently froze evolved cultures at -80 °C (transfer 8, 13, 22, 30, and 41), and periodically checked for contamination by plating a small aliquot of the culture on solid agar plates to observe colony morphology. Populations of strain RN were initially not included in the evolution experiment, and were evolved later along with large population size replicates (see below) with 36-hour transfers.

Populations of strains AF and AC (with nearly zero initial growth rate in methylamine) either went extinct or were rapidly overtaken by a contaminant that could easily outgrow the mutants. Therefore, we evolved these populations in larger population sizes in 50 ml Erlenmeyer flasks containing 10 mL media, and allowed them to gradually adapt to methylamine as follows. We inoculated 156.25 µL of each succinate-grown culture in 10 mL Hypho (1:64 dilution) containing 7.5 mM methylamine and 1.75 mM succinate (half the normal concentration of each carbon source). Four days later, we initiated the first transfer, inoculating 625 µL of culture in 10 mL Hypho broth (1:17 dilution) containing 15 mM methylamine as the sole carbon source. After four transfers, we increased dilution to 33 fold. After one more transfer, we increased dilution to 64 fold. In the following transfer, cultures were growing rapidly enough that they were moved to a 48-hour transfer cycle. Thus, by the end of the experiment AC and AF lines were evolving under similar conditions as the rest of the mutants.

Large population size: We allowed four replicate populations of each mutant to evolve in 30 ml media in 100 ml flasks maintained at 30 °C in a shaking incubator at 200 rpm. Each replicate population was initiated as above. Every 36 hours, we transferred 968 µL growing culture to 30 ml fresh Hypho media (1:32 dilution) containing 15 mM methylamine as the sole (selective) carbon source. For the wild type, this protocol allows approximately 5 generations per transfer. Every third transfer, we froze evolved cultures at -80 °C and checked for contamination by plating a small aliquot of the culture on solid agar

plates to observe colony morphology. One replicate population of AR was found to be contaminated with another bacterial species, and was discarded.

Populations of strains AF and AC (with nearly zero initial growth rate in methylamine) went nearly extinct. Therefore, we allowed these populations to gradually adapt to methylamine. We inoculated 968 µL of each succinate-grown culture in 30 mL Hypho (2:64 dilution) containing 3.75 mM methylamine and 2.625 mM succinate (1/4 of the normal methylamine concentration and 3/4 of the normal succinate concentration). After three transfers in this medium, the populations were evolved in Hypho medium with 7.5 mM methylamine supplemented with 1.75 mM succinate for another three transfers (half of the usual concentration of each carbon source). This was followed by six transfers in medium containing 15 mM methylamine as the sole carbon source. Sanger sequencing of isolates from the final frozen culture showed that one AF population was contaminated with the AC strain, so it was removed from subsequent analysis.

Calculating effective population size (N_e)

Four strains (AR, CO, VA and RN) had very similar initial growth rates, and hence similar expected N_e. Therefore, we only estimated starting effective population size for strains AF, AC, CO and WT. We allowed three replicate populations of each ancestor to grow under the same conditions used at the beginning of each experimental evolution line. Immediately after inoculating fresh media, we spread serial dilutions of a 50 µL aliquot on nutrient agar plates. After 5 days of incubation at 30 °C, we counted the number of colonies to estimate the initial population size (N_0). We similarly counted the final population size (N_f) at the end of the growth cycle, at 36 and 48 hours post-inoculation for large and small populations respectively. For each strain, we calculated the effective population size as: $N_e = g*N_0$, where $g = \log_2 (N_f/N_0)$.

Quantifying growth rate

We quantified growth rate of various strains as described earlier (Agashe et al. 2013). Briefly, we inoculated each strain from -80 °C freezer stocks into 5 mL Hypho-succinate medium in 50 mL culture tubes and incubated them at 30 °C, shaking at 200 rpm for 48 hours. We inoculated 10 µL of this culture in wells of a 48 well microplate containing 630 µL of Hypho + Succinate medium and incubated the plate at 30 °C at 200 rpm for 48 hours. We inoculated 10 µL culture from each well into a well of a fresh 48-well plate containing 630 µL of Hypho + 15mM methylamine. The plate was incubated at 30 °C at 170 rpm for 48 hours in a Tecan Infinite Series Multi-mode Reader (TECAN, Austria) and measured optical density at 15 minute intervals at 600 nm. Each of three independent biological replicates were distributed

in three technical replicates (wells). We calculated the maximum growth rate using the Curve Fitter software (Delaney, Kaczmarek, et al. 2013; Delaney, Rojas Echenique, et al. 2013).

Identifying *fae*-associated mutations in evolved lines

For each evolved population, we plated out an aliquot of final frozen stocks on Hypho-succinate agar plates and randomly chose five (large populations) or ten colonies (small populations) from each plate. We allowed each colony to grow in Hypho-succinate broth for 48 hours, extracted DNA from these cultures, and amplified and sequenced *fae* as well as the upstream and downstream regions, including the upstream gene *orf7* (**Table S3,** set "sequencing"). We sequenced the *fae* gene from each original clone used to initiate the evolution lines to confirm that the evolved allele was absent in the original clone.

Quantifying *fae* expression

Levels of *fae* mRNA were quantified for all strains as described in (Agashe et al. 2013) with the following modifications. We inoculated -80 °C freezer stocks of each strain in 5 mL Hypho + succinate medium in 50 mL culture tubes. After 2 days, we transferred 620 µL of this culture to 100 mL flasks containing 40 mL fresh Hypho + succinate medium. After 24 hours of growth in a shaking incubator, we added 600 µL of 1 M methylamine (15 mM final concentration) to each flask to induce *fae* expression. Two hours after induction, we split the grown culture into two 20 mL fractions for protein and mRNA quantification. We harvested cells by centrifugation (5000 rpm at 4 °C for 10-15 minutes) and immediately froze the pellet in -80 °C. To extract total RNA from the cells, we crushed pellets using liquid nitrogen (to lyse the cells), used the Trizol RNA extraction protocol (TRIzol, Ambion) and quantified total RNA (Nanodrop, Thermo Scientific). We made cDNA from each RNA sample (QuantiTect Reverse Transcription Kit, Qiagen). We set up quantitative PCR reactions with Maxima SYBR Green (Thermo scientific) in 384 well plates (Microamp endura plate optical, Applied Biosystems). We used specific primers to quantify *fae* mRNA for each mutant ((Agashe et al. 2013), Table S7).

Quantifying FAE protein

We quantified FAE protein levels from each strain as described in (Agashe et al. 2013) with the following modifications. Cells were grown and harvested as described above. We freeze-thawed cell pellets in liquid nitrogen 2-3 times for cell lysis, and extracted soluble proteins in the appropriate buffer as described previously. We quantified total protein in the cell lysate using the bicinchonic acid assay (BCA assay kit, Thermo-Fisher). We denatured proteins at 99 °C for 20 minutes with appropriate amounts of 6x SPLB (soluble protein lysis buffer). We loaded approximately equal amounts of total protein in each lane of a 12% 29:1 Acrylamide-Bisacrylamide SDS-PAGE gel, including one lane containing 6 µL of a

protein molecular weight marker (Spectra broad range protein ladder, Thermo-Fisher). We ran the gel for 1.5 hours at 150 Volts (Bio-Rad Mini-Protean Tetra-Cell). We transferred the proteins onto a PVDF membrane (Immobilon-P, Merck-Millipore) at 100 V for 1 hour (Bio-Rad Mini-Protean Wet Transfer System). After transfer, we cut the blot in half at the 50 kDa marker band. The top half of the blot (high molecular weight) was Coomassie stained and the bottom half of the blot (low molecular weight) was blocked as described previously. We probed the bottom half with anti-FLAG antibody raised in mouse (Sigma) at a dilution of 1:10000 in 5% non-fat dry milk in TBS with 0.1% Tween-20, and incubated the blot at 4 °C on a rocker overnight, followed by three 20 minute washes with TBS with 0.1% Tween-20. We incubated the blot in secondary HRP-conjugated anti-mouse IgG antibody raised in rabbit (Sigma) (diluted 1:5000 in 5% non-fat dry milk in TBS+0.1% Tween-20) for 1 hour on a rocker at room temperature, followed by three 20-minute washes with TBS+0.1% Tween-20. The blot was developed using SuperSignal West-Femto Chemiluminescence substrate (Thermo) and imaged in the ImageQuant LAS 4000 imager (GE Biosciences). Chemiluminescence signal for FAE in each lane was quantified using ImageJ and normalized to the intensity of Coomassie stained bands in the top half of the blot in the corresponding lane (also quantified using ImageJ).

Quantifying FAE enzyme activity

FAE catalyzes the condensation of formaldehyde with tetrahydromethanopterin (H_4MPT) to form methylene-H_4MPT (Vorholt et al. 2000). FAE activity can be tracked using a coupled assay with MtdB ($NAD(P)^+$-dependent methylene-H_4MPT dehydrogenase), which generates methenyl-H_4MPT and NAD(P)H (Vorholt et al. 2000). We purified His-tagged MtdB (MtdB-His6) and H_4MPT as described previously (Agashe et al. 2013). We used 190 g of cell paste from wild type *M. extorquens* AM1 grown on methanol (125 mM) to extract the cofactor (H_4MPT). We performed the activity assay as described earlier (Vorholt et al. 2000) with the following differences: we used 50 mM Potassium Phosphate buffer pH 7.8, purified MtdB-His6 (200 μM, 50 μL) and co-substrate $NADP^+$ (125 μM). The total volume of the reaction was 0.5 mL. We monitored enzyme activity (NADPH production) at 340 nm after addition of formaldehyde (2 mM), calculating the rate of the reaction using the slope of the linear increase in absorbance over time. One unit of activity was defined as the rate of formation of methenyl-H_4MPT per mg total protein in cell lysate (μmol/mg/min).

Estimating mRNA folding energy

We created all possible single variants of each ancestral *fae* allele *in silico* using an R (R Core Team 2015) script, mutating each of the first 50 nucleotides in turn. We measured the Minimum Folding Energy (MFE) of each mutant, the *fae* ancestral and evolved alleles, and alleles carrying the N-terminal insert in

sliding windows of 50 nt using the RNAslider program with default settings (Horesh et al. 2009). The program uses a modified RNAfold algorithm from the Vienna RNA package (Lorenz et al. 2011) to reduce time required for calculations. The MFE of each window was assigned to the first base of the window (e.g. the MFE of bases 0 to 50 was assigned to the MFE at position 0). We then determined the MFE for all alleles at 30 °C using the RNAfold program (Lorenz et al. 2011) and calculated the difference between the MFE of each modified allele and the respective ancestral allele.

In the model used to calculate minimum folding energy by the program RNAfold, temperature and dangling energies are the parameters that affect MFE values the most. As explained above, we calculated MFE values at 30° C, which is the optimum growth temperature of *M. extorquens*. To test the impact of temperature on our results, we calculated MFE values for all alleles at different temperatures ranging from 10° C to 70° C and determined their correlation with values at 30° C. We found that for temperatures between 25° C and 40° C, the correlation coefficient (rho) was > 0.95 (Fig S11) indicating that MFE values are robust to changes in folding temperature in the model. Similarly, we calculated MFE values by excluding dangling energies and found a strong correlation with MFE values with or without dangling energies (Fig S12).

Estimating anti-SD binding affinity

We calculated the anti-Shine-Dalgarno (SD) binding affinity of all hexamers occurring in the first 50 bp of ancestral and modified *fae* alleles, measured as the ΔG value of binding between each hexamer and the anti-SD sequence (Li et al. 2012). We used binding affinities for all possible 4096 hexamers calculated for *Escherichia coli* (Li et al. 2012). The SD sequence for *fae* is empirically unknown and thus, we used information for *Methylobacterium extorquens* AM1 from the gene finding algorithm, PRODIGAL (ftp://ftp.ncbi.nih.gov/genomes/Bacteria/Methylobacterium_extorquens_AM1_uid57605/NC_012808.Prodigal-2.50). The predicted ribosome binding motif for *fae* was GGAG/GAGG. Hexamers that included these motifs were included in the top 4% (top 161 hexamers) of binding affinity values in the above distribution for *Escherichia coli*. Thus, we could reliably use the same distribution for calculating anti-SD binding affinities of the *fae* alleles. A single base change can potentially alter anti-SD affinity in six frames (generating six new hexamers), so we ascribed the sum of each new hexamer's affinity to the point mutation. Hence, we calculated change in anti-SD binding affinity of each mutated *fae* variant as the difference between the total anti-SD binding affinity of the variant and its respective ancestral allele.

To confirm that our results were robust to variation in the exact anti-SD sequence (which is empirically uncharacterized for *M.* extorquens), we recalculated hexamer affinity values using the 3' tail of the 16S rRNA of *M. extorquens* AM1. We used these values to predict the effect of all possible point mutations on anti-SD affinity of all alleles. We still found no correlation between Δanti-SD affinity and fitness of

the alleles that we generated by site-directed mutagenesis. We also tested whether our results were robust to the cumulative anti-SD affinity as calculated above in all possible frames (for a given point mutation, six hexamers are changed in the allelic sequence). Instead of considering the anti-SD affinity change in all six hexamers, we only considered the highest Δanti-SD affinity among the six changed hexamers. We found that the lack of correlation between Δanti-SD affinity and fitness persisted (Fig S13). Note that in some cases a single point mutation alleviated the anti-SD affinity of one of the six changed hexamers, whereas it increased the anti-SD affinity of an adjacent hexamer. Thus, cumulative calculation maybe a better indicator for the potential effect of a point mutation on the fitness of the alleles.

SUPPLEMENTARY FIGURES

Figure S1: Time course of growth rate of evolving populations. Time series of growth rate of replicate populations as a function of approximate number of generations. Solid lines = small populations; dashed lines = large populations (samples from large evolving populations were frozen more frequently); horizontal dotted lines = WT growth rate; *n* = 3 technical replicates per population per time point.

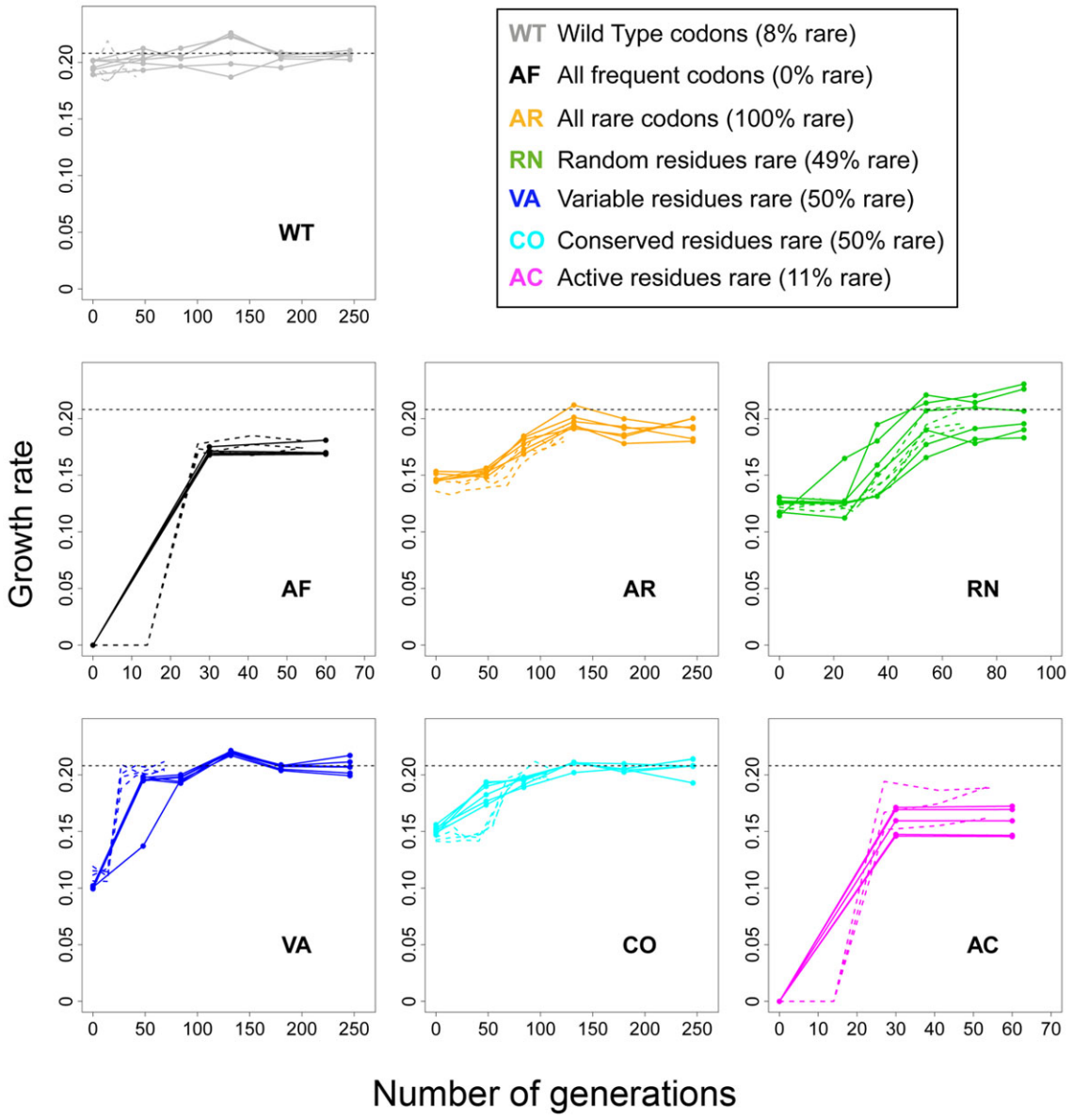


Figure S2: Comparison of ancestral and evolved phenotypes. Mean (±SEM) values for (A) growth rate (B) *fae* mRNA and (C) FAE protein of ancestral strains ("anc") compared to evolved isolates carrying an *fae*-associated SNP ("evo"). mRNA and protein values are shown relative to wild type (e.g. $mRNA_{anc}/mRNA_{WT}$). Asterisks indicate evolved isolates with significantly different values than their ancestor ($p < 0.05$ after correcting for multiple comparisons using the Benjamini-Hochberg method that controls the false discovery rate). Strains are colored as in Fig S1.

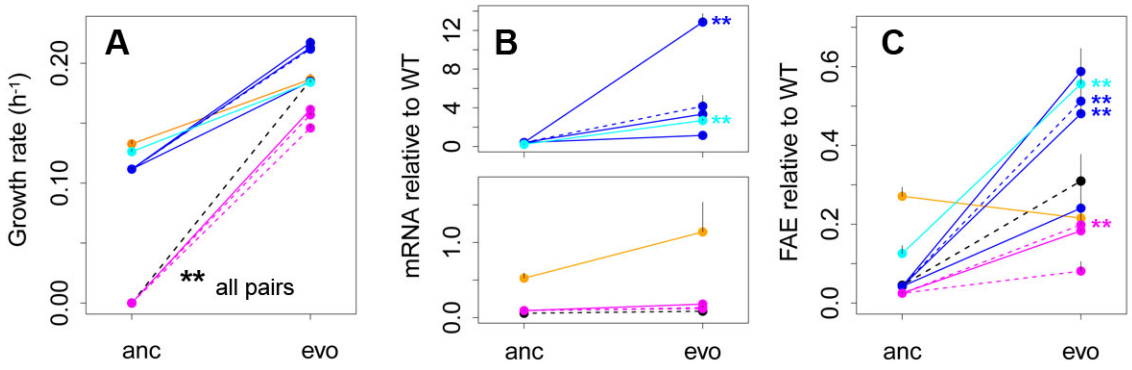


Figure S3: Time course of allele frequencies. Time series of allele frequencies in individual populations as a function of approximate number of generations. Data were available only for replicates with small population size. For intermediate time steps, n = 3 isolated colonies per population; for the final time step, n = 10 isolated colonies per population. Dashed lines = polymorphism, dotted lines = no mutation (i.e. ancestor).

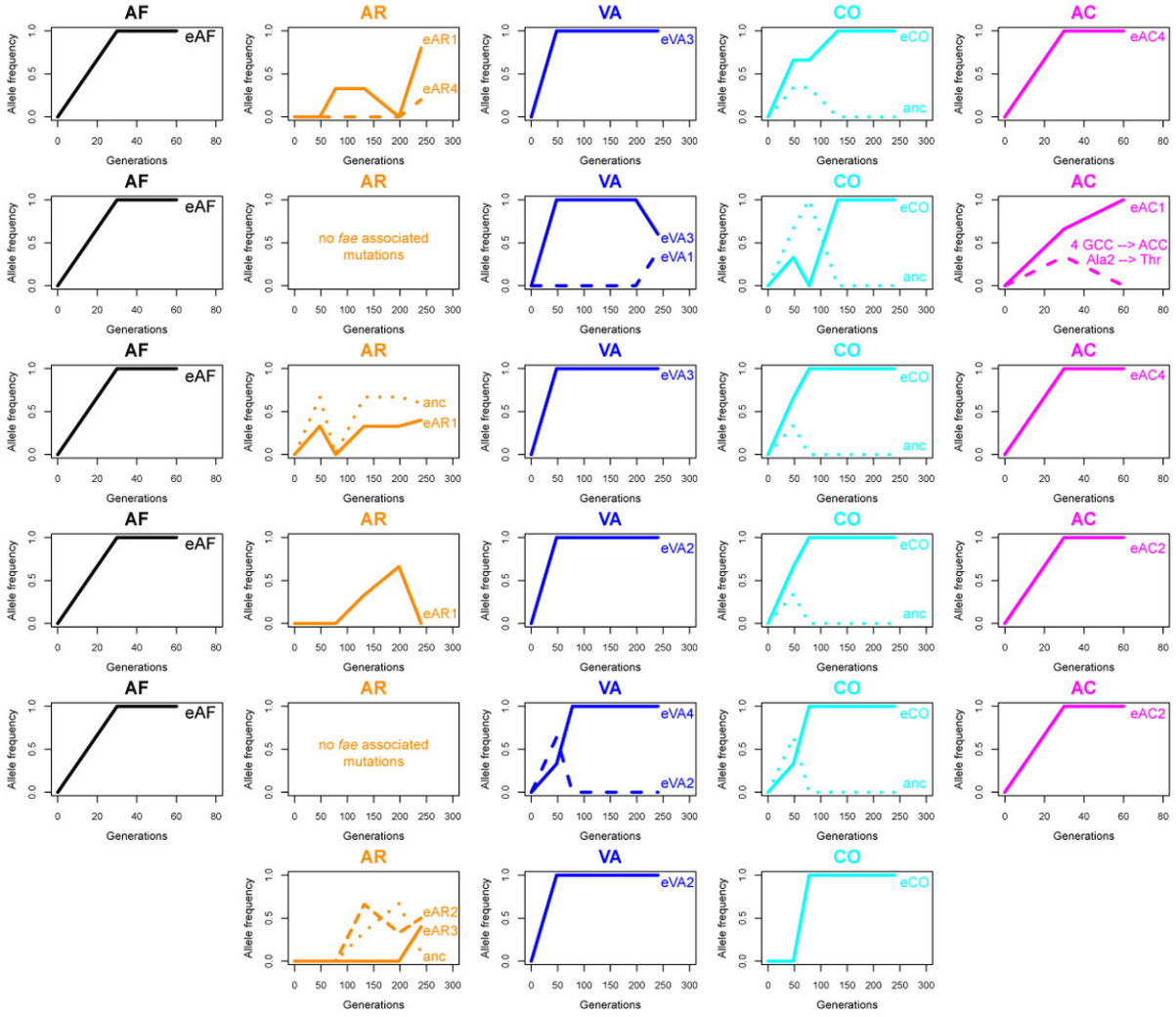


Figure S4: Comparison of phenotypic impacts of SNPs and genomic background. Mean (+SEM) values for (A) growth rate (B) *fae* mRNA and (C) FAE protein of ancestral strains ("anc") compared to evolved isolates carrying the ancestral *fae* allele ("bkg", for genomic background mutations), ancestral strains carrying an evolved *fae*-associated SNP ("snp"), and evolved isolates carrying an *fae*-associated SNP ("evo"). mRNA and protein values are shown relative to wild type (e.g. $mRNA_{anc}/mRNA_{WT}$). Strains are colored as in Fig S1. Significantly different pairwise comparisons are marked with an asterisk and the corresponding SNP is labeled (*p = 0.05, **p < 0.05,,after correcting for multiple comparisons using the Benjamini-Hochberg method). E.g. in panel A, an evolved isolate carrying the eVA3 SNP also had a genomic background mutation that significantly increased growth rate compared to ancestral VA; in all other cases, background mutations did not affect growth rate. In contrast, all evolved SNPs had a different growth rate compared to their genomic background (*all pairs), and in two cases (eVA3 and eAR1), the fitness effect of the SNP was different than that of the evolved isolate in which it was found.

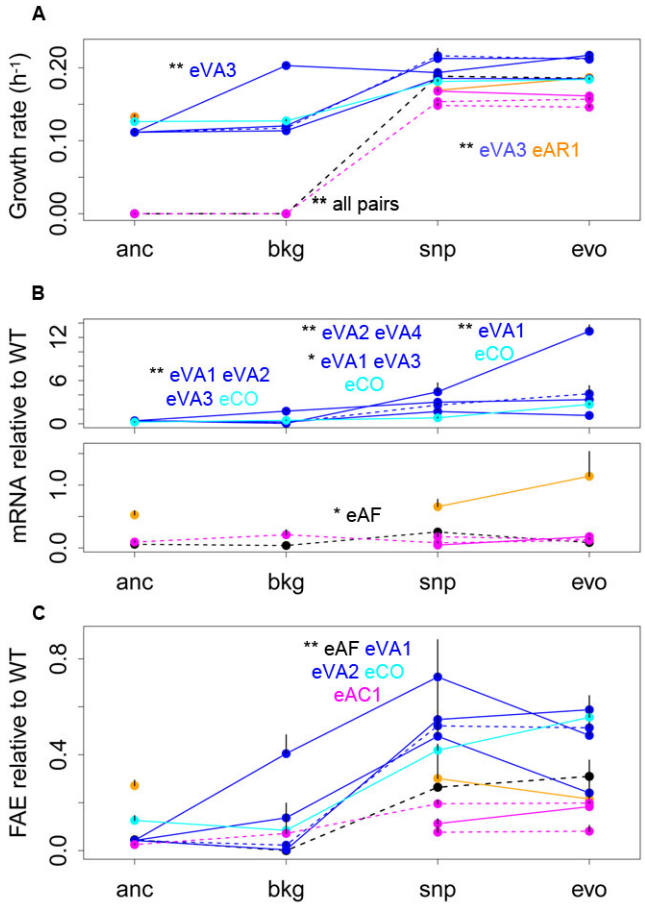


Figure S5: Comparison of phenotypic effects of synonymous and nonsynonymous SNPs. In each case, an ANOVA indicates that the impact of nonsynonymous (NS) and synonymous (S) mutations is not significantly different ($p > 0.2$). Data shown in each panel were calculated as the effect of SNP relative to ancestor, after normalizing to the mean WT value in each case to allow comparison.

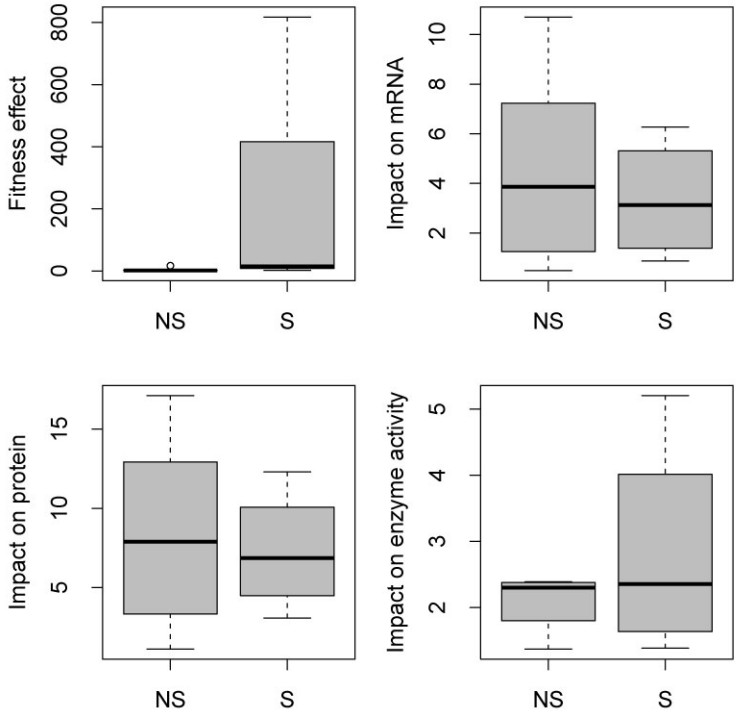


Figure S6: Association between fitness and N-terminal sequence features. Growth rate of *fae* variants as a function of (**A**) the predicted 5' MFE and (**B**) predicted number of hexamers with high anti-SD affinity (ΔG < -4 kcal/mol). Points are colored by strain as in Figure S1. For MFE calculations we used 100 bp long sequences (-50 to +50 bp relative to the *fae* start site). For anti-SD affinity calculations, we used 55 bp long sequences (-5 to +50 bp relative to the *fae* start site).

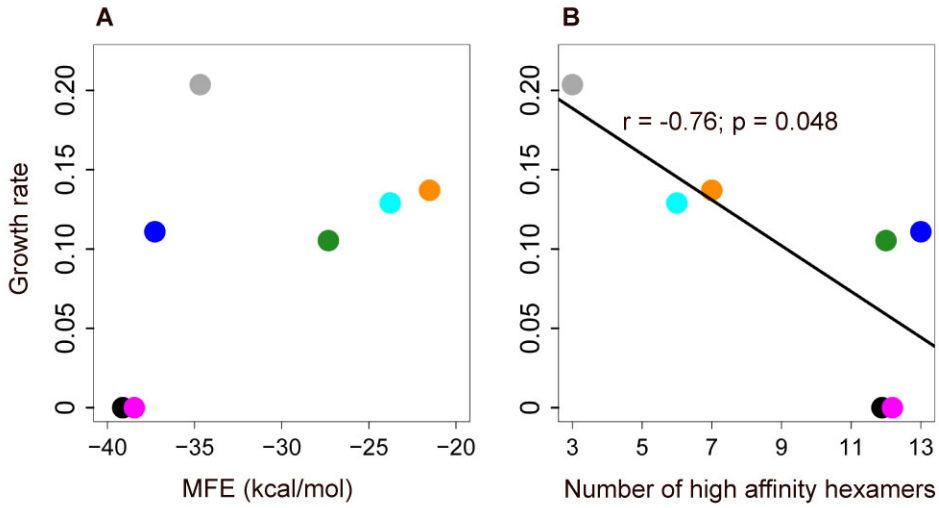


Figure S7: Predicted effect of *fae*-associated SNPs on mean folding energy (MFE) and anti-SD binding affinity of AR, CO and RN strains (ancestor – mutant). Dashed lines indicate no effect on anti-SD affinity and MFE. Grey points = all possible mutations in the first 50 bases of the coding sequence; red triangles = evolved mutations in each strain.

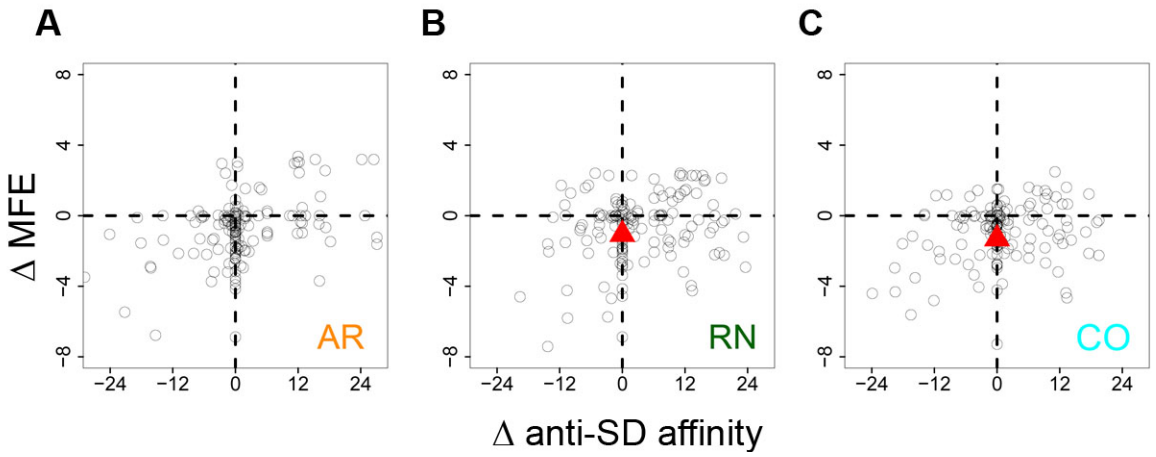


Figure S8: Distribution of effects of all possible mutations on (A) anti-SD affinity and (B) MFE of each ancestral *fae* allele. Evolved alleles are marked by vertical dashed lines.

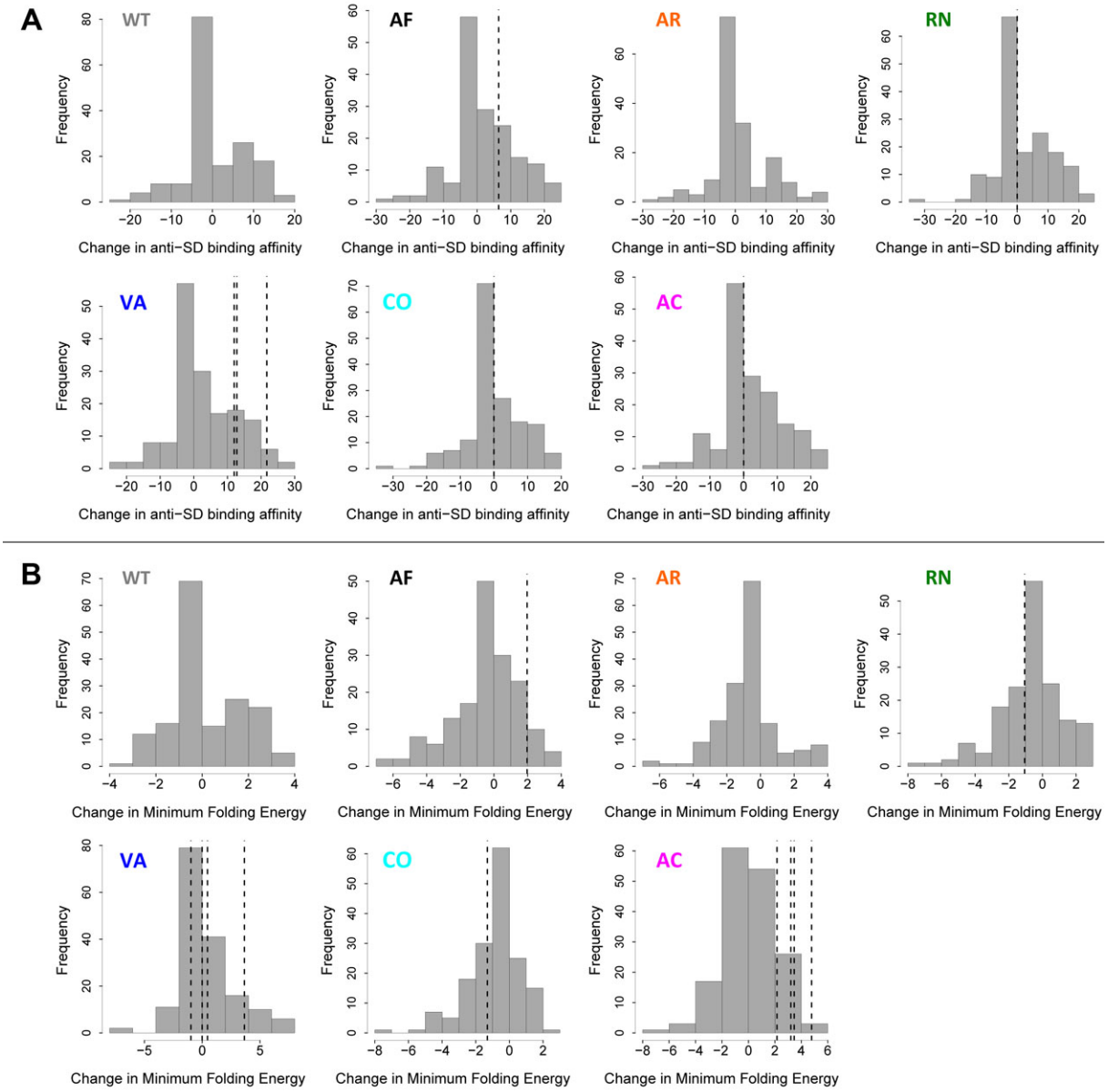


Figure S9: Change in fitness as a function of (A) change in anti-SD affinity or (B) MFE. Data are shown for strains carrying plasmid-borne *fae* alleles with point mutations. Triangles = evolved *fae* alleles; filled circles = engineered mutations (see Fig 5, I-J). Alleles are colored as indicated in Figure S1.

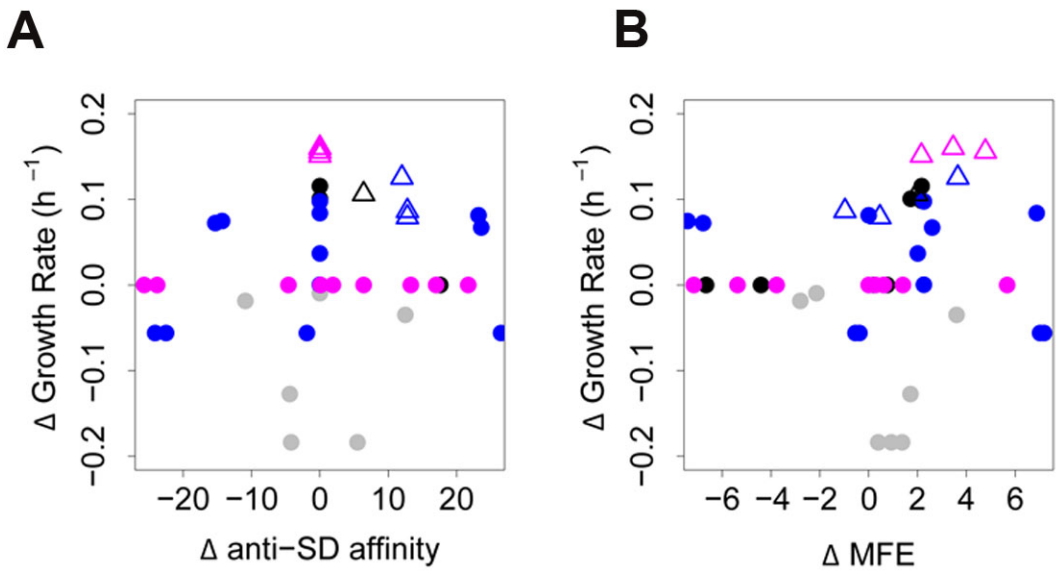


Figure S10: Comparison of fitness effect of Δ*fae* strains carrying plasmid-borne synonymous and nonsynonymous SNPs in WT, AF, AC and VA alleles. A Wilcoxon rank sum test indicates that the fitness effect of nonsynonymous (ns) and synonymous (syn) point mutations (estimated as growth rate relative to ancestor) is not significantly different (p = 0.095).

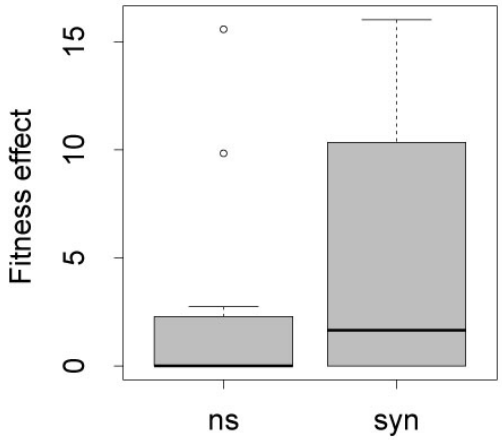


Fig S11: MFE as a function of folding temperature. The correlation coefficient (Spearman's rho) between MFE calculations at 30°C and at various other temperatures is shown as a function of the alternative temperature, for each ancestral *fae* allele. Red vertical lines indicate the growth temperature of *M. extorquens,* which was used for MFE calculations reported and analyzed in this study.

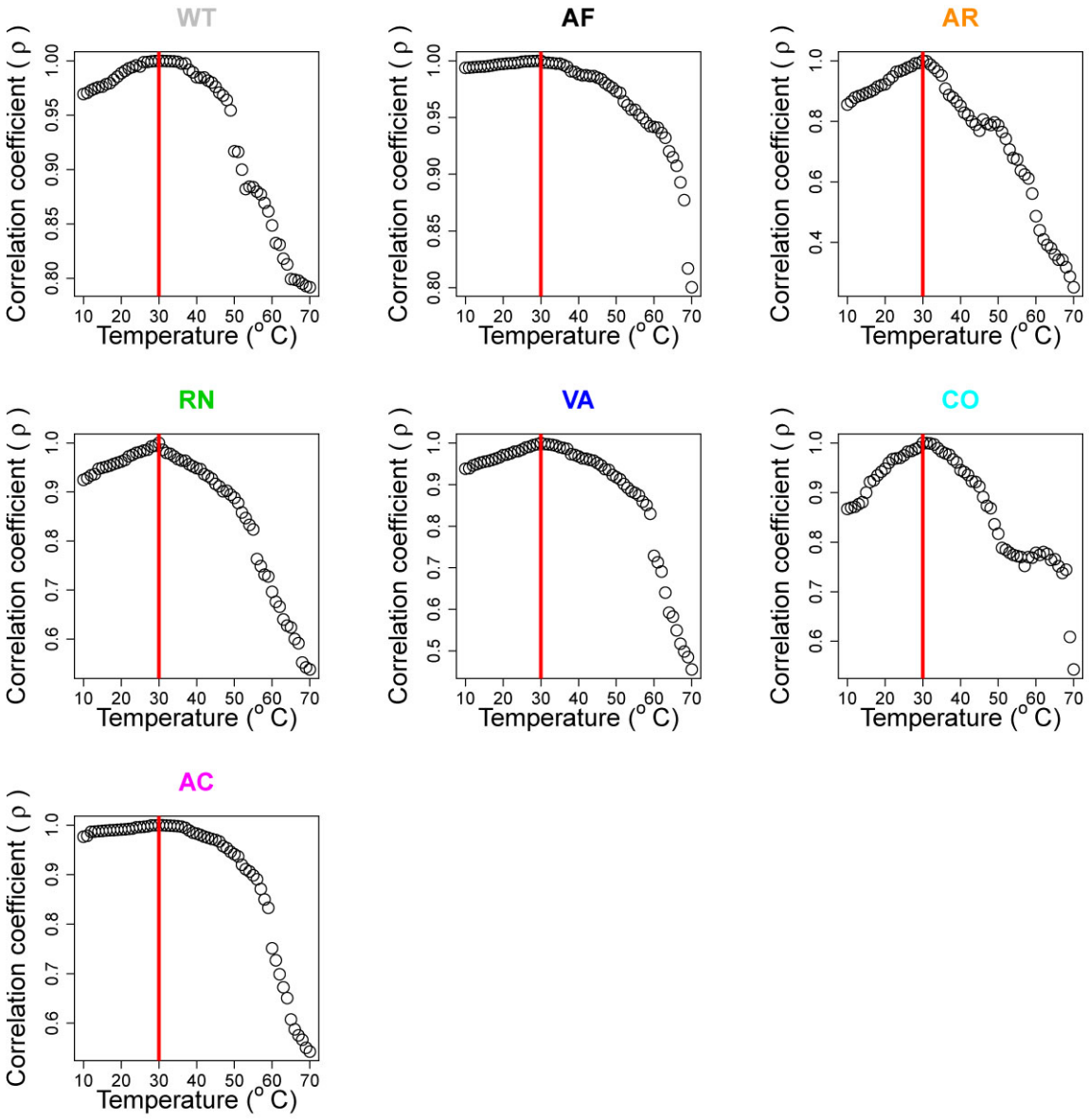


Fig S12: MFE calculating with or without free energy of dangling ends. Plots show the correlation between MFE calculated with vs. without contribution from dangling end free energies, for each ancestral *fae* allele and all possible point mutations 50 bp upstream and downstream of the start site for each allele. Dashed lines indicate equal MFE. Spearman's rho for the correlation is indicated in each panel.

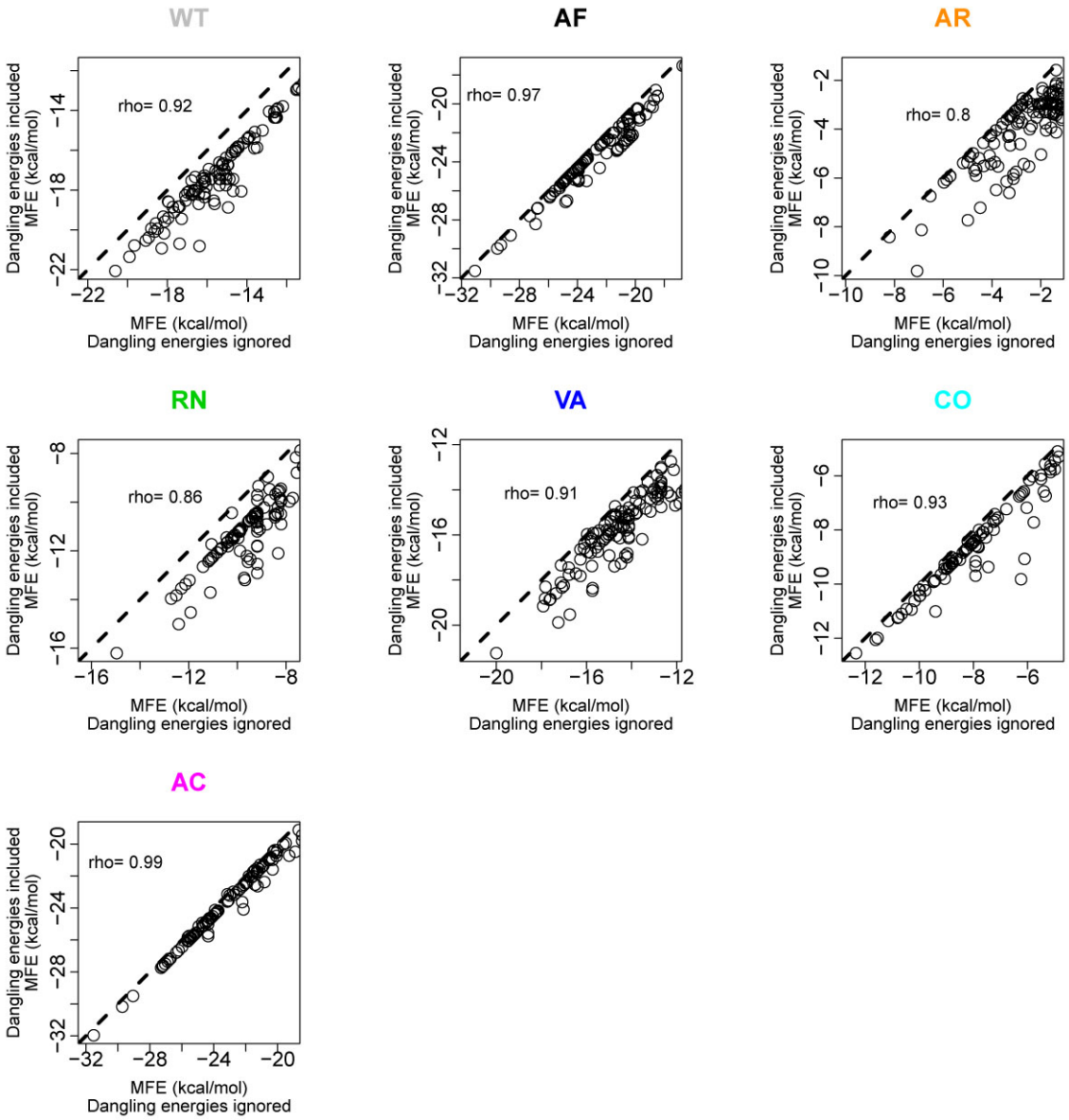
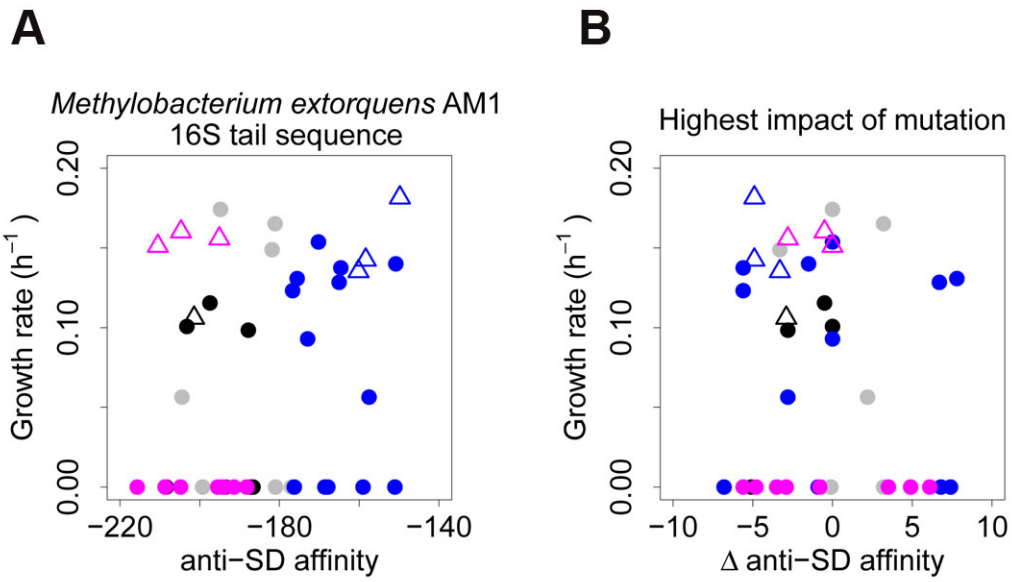


Fig S13: Impact of using alternative methods to calculate anti-SD affinity. (A) Growth rate of strains carrying plasmid-borne *fae* alleles (see Fig 5) as a function of the anti-SD affinity of the respective alleles. Affinity was calculated using the *M. extorquens* AM1 16S rRNA tail sequence (13 nts as described by(Nakagawa et al. 2010)) (B) Growth rate as a function of the maximum impact of a point mutation on the anti-SD affinity of the ancestral allele, in any frame. Triangles = evolved *fae* alleles; filled circles = engineered mutations. Alleles are colored as indicated in Figure S1.



SUPPLEMENTARY REFERENCES

Agashe D, Martinez-Gomez NC, Drummond DA, Marx CJ. 2013. Good codons, bad transcript: Large reductions in gene expression and fitness arising from synonymous mutations in a key enzyme. Molecular Biology and Evolution 30:549–560.

Chou H-H, Marx CJ. 2012. Optimization of gene expression through divergent mutational paths. CellReports 1:133–140.

Chubiz LM, Purswani J, Carroll SM, Marx CJ. 2013. A novel pair of inducible expression vectors for use in *Methylobacterium extorquens*. BMC Res Notes 6:183.

Delaney NF, Kaczmarek ME, Ward LM, Swanson PK, Lee M-C, Marx CJ. 2013. Development of an optimized medium, strain and high-throughput culturing methods for *Methylobacterium extorquens*. PLoS ONE 8:e62957.

Delaney NF, Rojas Echenique JI, Marx CJ. 2013. Clarity: An open-source manager for laboratory automation. Journal of Laboratory Automation 18:171–177.

Horesh Y, Wexler Y, Lebenthal I, Ziv-Ukelson M, Unger R. 2009. RNAslider: a faster engine for consecutive windows folding and its application to the analysis of genomic folding asymmetry. BMC Bioinformatics 10:76.

Li G-W, Oh E, Weissman JS. 2012. The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. Nature 484:538–541.

Lorenz R, Bernhart SH, Hoener Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. 2011. ViennaRNA Package 2.0. Algorithms for Molecular Biology 6:26.

Marx CJ, Lidstrom ME. 2002. Broad-host-range *cre-lox* system for antibiotic marker recycling in gram-negative bacteria. BioTechniques 33:1062–1067.

Marx CJ. 2008. Development of a broad-host-range *sacB*-based vector for unmarked allelic exchange. BMC Res Notes 1:1.

Nakagawa S, Niimura Y, Miura KI, Gojobori T. 2010. Dynamic evolution of translation initiation mechanisms in prokaryotes. Proceedings of the National Academy of Sciences 107:6382–6387.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing [Internet]. Available from: http://www.R-project.org/

Vorholt JA, Marx CJ, Lidstrom ME, Thauer RK. 2000. Novel formaldehyde-activating enzyme in *Methylobacterium extorquens* AM1 required for growth on methanol. Journal of Bacteriology 182:6645–6650.